AMENDMENTS TO THE SPECIFICATION

On page 3, please replace the paragraph beginning on line 7 and ending on line 15 with the following amended paragraph:

In still other embodiments, the nucleic acid sequence encoding a polypeptide at least 40% identical to SEQ ID NO: 1 is selected from the group consisting of SEQ ID NOs: 1-4, 16-21, 33-39, 49-52, and 56. In further embodiments, the nucleic acid sequence is selected from the group consisting of SEQ ID NOs: 5-9, 22-27, 40-48, 53-55, and 57~~ and 58~~. Accordingly, in some embodiments the present invention provides expression vectors comprising nucleic acid sequences at least 40% identical to any one of SEQ ID NOs: 5-9, 22-27, 40-48, 53-55, and 57~~ and 58~~. In further embodiments, the nucleic acid sequence is at least 40%, 60%, 70%, 80%, 90%, 95% (or more) identical to any of SEQ ID NOs: 5-9, 22-27, 40-48, 53-55, and 57~~ and 58~~.

On page 7, please replace the paragraph beginning on line 8 and ending on line 16 with the following amended paragraph:

Fig. 2. shows exemplary embodiments which demonstrates *(A)* positional cloning of the *LUT1* locus showing recombinants as indicated for specific SSLP markers across the interval and the position of chloroplast-targeted proteins are indicated by dashed arrows, *(B)* overview of the intron-exon organization of *LUT1* and the locations of the *lut1-1* and *lut1-3* mutations, and *(C)* Deduced amino acid sequence of LUT1 (SEQ ID NO:4). The cleavage site of the putative chloroplast targeting sequence is indicated by an arrow and the single predicted transmembrane domain is shaded in black. The conserved cytochrome P450 molecular oxygen binding pocket and the cysteine motif are indicated by single and double underlines, respectively, and the conserved Thr by an asterisk.

On page 8, please replace the paragraph beginning on line 16 and ending on line 19 with the following amended paragraph:

Fig. 10. (SEQ ID NOS:60-74) shows exemplary embodiments that demonstrate sequence similarities of CYP97A and CYP97C sequences. A rooted neighbor-joining tree was

constructed using CYP97B from *Arabidopsis thaliana* as an outgroup. Bootstrap values are indicated adjacent to the branches.

On page 9, please replace the paragraph beginning on line 21 and ending on line 24 with the following amended paragraph:

Fig. 21. ~~SEQ ID NO: 8: shows an upstream region nucleotide sequence of leaky mutant *Arabidopsis thaliana* LUT1-2 (*lut1-2*) (Brassicaceae; thale cress). SEQ ID NO: 9: shows an a cDNA nucleotide sequence of knockout mutant *Arabidopsis thaliana* LUT1-3 sequence (*lut1-3*) (Brassicaceae; thale cress).~~ shows exemplary embodiments which demonstrate (A) a leaky mutant resulting from a rearrangement in the upstream region in *Arabidopsis thaliana (lut1-2)* (Brassicaceae; thale cress) and (B) shows a knockout mutant in *Arabidopsis thaliana* resulting from a T-DNA insertion in the sixth intron (*lut1-3*) (Brassicaceae; thale cress).

On page 13, please replace the paragraph beginning on line 4 and ending on line 5 with the following amended paragraph:

Fig. 32. ~~SEQ ID NO: 58: shows a knockout mutant *Arabidopsis thaliana* CYP97A3 (Brassicaceae; thale cress).~~ shows an exemplary embodiment which demonstrates a single knockout mutant of *Arabidopsis thaliana* CYP97A3 (Brassicaceae; thale cress) (SALK_116660, herein incorporated by reference).

On page 40, please replace the paragraph beginning on line 9 and ending on line 18 with the following amended paragraph:

The terms "*lut1* gene" or "*lut1*" or "*lutein* gene" refer to a plant gene in which a knock-out mutation results in partial or complete loss of lutein, or alteration of carotenoid ratios, in a genetic background where the wild type or non-mutant phenotype (containing the wild type *LUT1* gene) produces lutein (as demonstrated in Figs. 1, 3 and 4). The terms "*lut1* gene," "*lut1-1*," "*lut1-2*" or "*lut1-3*," and the like, refer to specific *LUT1* alleles e.g., SEQ ID NOs:~~6-10~~ 6-7, 10 and 23-28. The present invention identifies *lut1* genes that are referred to by number, for example, *lut1, lut1-1, lut1-2*, and *lut1-3*. The present invention identifies *lut1* polypeptides encoded by *lut1* genes; these polypeptides are referred to by number, for

example, LUT1, lut1-1, lut1-2 and lut1-3, e.g., SEQ ID NOs:~~4, 7-9 and 58~~ 4 and 7 and Figs. 2B and 2C.

On page 41, please replace the paragraph beginning on line 4 and ending on line 26 with the following amended paragraph:

The present invention is not limited to the use of any particular homolog or variant or mutant of LUT 1 protein or *lut1* gene. Indeed, in some embodiments a variety of LUT 1 protein or *lut1* genes, variants and mutants may be used so long as they retain at least some of the activity of the corresponding wild-type protein. In some embodiments, a variety of LUT 1 protein or *lut1* genes, variants and mutants may be used so long as they increase the activity of the corresponding wild-type protein. In particular, it is contemplated that proteins encoded by the nucleic acids of SEQ ID NOs:~~5-9~~ 5-7, 22-27, 40-48, and 53-56~~, and 58~~ find use in the present invention. In particular, it is contemplated that nucleic acids encoding proteins that comprise polypeptides at least 40% identical to SEQ ID NO: 1 and the corresponding encoded proteins find use in the present invention. Accordingly in some embodiments, the percent identity is at least 50%, 60%, 70%, 80%, 90%, 95% (or more). In still other embodiments, the nucleic acid sequence further comprises a sequence encoding a cytochrome P450 molecular oxygen binding pocket conserved consensus amino acid motif corresponding to SEQ ID NO:12. In other embodiments, the nucleic acid sequence further comprises a sequence encoding a conserved transmembrane domain sequence corresponding to SEQ ID NO: 10. In further embodiments, the nucleic acid sequence further comprises a sequence encoding a conserved consensus cysteine motif in P450 molecules corresponding to SEQ ID NO: 14. In other embodiments, the nucleic acid sequence further comprises a sequence encoding a LUT1 conserved consensus cysteine amino acid motif corresponding to SEQ ID NO:15. In still further embodiments, the nucleic acid sequence further comprises a sequence encoding a conserved N-terminal transit peptide for chloroplast-targeting corresponding to SEQ ID NO:11.

On page 42, please replace the paragraph beginning on line 18 and ending on page 43, line 5 with the following amended paragraph:

The *LUT1* locus has previously been mapped to the bottom arm of chromosome 3 at 67 ± 3 cM (Tian, *et al. Plant Mol. Biol.* 47, 379-388 (2001), herein incorporated by reference). For fine mapping of the locus, 530 plants homozygous for the *lut1* mutation were identified from approximately 2,000 plants in a segregating F_2 mapping population. Using SSLP markers, *LUT1* was initially localized to an interval spanning two BAC clones (F8J2 and T4D2) and was further delineated to a 100 kb interval containing 30 predicted proteins (Fig. 2*A*). The term "BAC" and "bacterial artificial chromosome" refers to a vector carrying a genomic DNA insert, typically 100-200 kb. The term "SSLP" and "simple sequence length polymorphisms" refers to a unit sequence of DNA (2 to 4 bp) that is repeated multiple times in tandem wherein common examples of these in mammalian genomes include runs of dinucleotide or trinucleotide repeats (for example, CACACACACACACACACA (SEQ ID NO:59)." As with all other carotenoid biosynthetic enzymes, the *LUT1* gene product is predicted to be chloroplast-targeted and within the 100 kb interval containing *LUT1*, six proteins were predicted as being chloroplast-targeted by the TargetP prediction software (http://www.cbs.dtu.dk/services/TargetP). One of these chloroplast-targeted proteins, At3g53130, is a member of the cytochrome P450 monooxygenase family (CYP97C1). Cytochrome P450 monooxygenases are heme-binding proteins that insert a single oxygen atom into substrates, e.g. hydroxylation reactions, and therefore At3g53130 was considered to be a strong candidate for *LUT1*.

On page 46, please replace the paragraph beginning on line 8 and ending on line 17 with the following amended paragraph:

The deduced amino acid sequence of LUT1 contains several features characteristic of cytochrome P450 enzymes (Fig. 2C). Cytochrome P450 monooxygenases contain a consensus sequence of (A/G)GX(D/E)T(T/S) (SEQ ID NO:12) that forms a binding pocket for molecular oxygen with the invariant Thr residue playing a critical role in oxygen binding in both prokaryotic and eukaryotic cytochrome P450s (Chapple, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 49, 311-343 (1998, herein incorporated by reference). In the deduced LUT1 protein sequence, this oxygen-binding pocket is highly conserved (single underlined amino acids in Fig. 2C). The conserved sequence around the heme-binding cysteine residue for cytochrome

P450 type enzymes is FXXGXXXCXG (SEQ ID NO:14), and is also present in LUT1 (double underlined amino acids in Fig. 2*C*).

On page 48, please replace the paragraph beginning on line 23 and ending on page 49, line 2 with the following amended paragraph:

The present invention provides plant *LUT1* genes and proteins including their homologs, orthologs, paralogs, variants and mutants. The designation "LUT" refers to the phenotype exhibited by plants with a mutation in a *LUT1* gene (the mutant allele is termed *lut1*), where the mutant has lowered levels of lutein (also referred to as decreased ε-ring hydroxylase activity). In some embodiments of the present invention, isolated nucleic acid sequences comprising *LUT1* genes are provided. Mutations in these genes, which disrupt expression of the genes, result in altered carotenoid ratios and carotenoid phenotype. In some embodiments, isolated nucleic acid sequences comprising *lut1-1*, *lut1-2*, *lut1-3 or CYP97C or CYP97B* are provided. These sequences include sequences comprising *lut1* and *CYP97C* cDNA/genomic sequences (for example, as shown in Figs. 2B, 2C and Fig. 7; ~~SEQ ID NOs: 7-9 and 58)~~ SEQ ID NO:4.

On page 49, please replace the paragraph beginning on line 5 and ending on line 22 with the following amended paragraph:

The present invention provides nucleic acid sequences comprising additional CYP97 cytochrome P450 genes. For example, some embodiments of the present invention provide polynucleotide sequences that produce polypeptides that are homologous to at least one of SEQ ID NOs: 1-3. In some embodiments, the polypeptides are at least 40%, 60%, 70%, 80%, 90%, 95% (or more) identical to any of SEQ ID NOs: 1-4, 16-21, 33-39, 49-52 and 56. Other embodiments of the present invention provide sequences assembled through EST sequences that produce polypeptides at least 40% or more (e.g., 60%, 70%, 80%, 90%, 95%) identical to at least one of SEQ ID NOs: 11-14, 16-21, 33-39, 49-52 and 56. In other embodiments, the present invention provides nucleic acid sequences that hybridize under conditions ranging from low to high stringency to at least one of SEQ ID NOs: ~~5-9~~ 5-7, 22-27, 40-48, and 53-56, ~~and 58,~~ as long as the polynucleotide sequence capable of hybridizing to at least one of SEQ ID NOs: ~~5-9~~ 5-7, 22-27, 40-48, 53-55, and 57 ~~and 58~~

encodes a protein that retains a desired biological activity of a carotenoid hydroxylase protein; in some preferred embodiments, the hybridization conditions are high stringency. In preferred embodiments, hybridization conditions are based on the melting temperature (T_m) of the nucleic acid binding complex and confer a defined "stringency" as explained above (See *e.g.*, Wahl *et al.*, Meth. Enzymol., 152:399-407 (1987), incorporated herein by reference).

On page 50, please replace the paragraph beginning on line 3 and ending on line 13 with the following amended paragraph:

In other embodiments of the present invention, the polynucleotide sequence encoding a *CYP97* gene is extended utilizing the nucleotide sequences (*e.g.*, SEQ ID NOs: ~~5-9~~ 5-7, 22-27, 40-48, 53-55, and 57 ~~and 58~~) in various methods known in the art to detect upstream sequences such as promoters and regulatory elements. For example, it is contemplated that for *LUT1*, *lut1-1*, *lut1-2*, *lut1-3*, or related *CYP97* hydroxylases, the sequences upstream are identified from the Arabidopsis genomic database. For other *lut1* genes for which a database is available, the sequences upstream of the identified *lut1* genes can also be identified. An example of an allele for an upstream region is shown is described herein as *lut1-2* ~~(SEQ ID NO: 8)~~. For other *lut1* and *CYP97* genes for which a public genomic database is not available, or not complete, it is contemplated that polymerase chain reaction (PCR) finds use in the present invention.

On page 52, please replace the paragraph beginning on line 25 and ending on page 53, line 13 with the following amended paragraph:

It is contemplated that is possible to modify the structure of a peptide having an activity (*e.g.*, such as a hydroxylase activity), for such purposes as increasing synthetic activity or altering the affinity of the LUT1 protein for a binding partner or a kinetic activity. Such modified peptides are considered functional equivalents of peptides having an activity of a LUT1 activity as defined herein. A modified peptide can be produced in which the nucleotide sequence encoding the polypeptide has been altered, such as by substitution, deletion, or addition. In some preferred embodiments of the present invention, the alteration increases or decreases the effectiveness of the *lut1* gene product to exhibit a phenotype caused by altered carotenoid production. In other words, construct "X" can be evaluated in order to

determine whether it is a member of the genus of modified or variant *lut1* genes of the present invention as defined functionally, rather than structurally. Accordingly, in some embodiments the present invention provides nucleic acids comprising a *lut1* or *CYP97* sequence that complement the coding regions of any of SEQ ID NOs: ~~5-9~~ 5-7, 22-27, 40-48, 53-55, and 57 ~~and 58~~, as well as the polypeptides encoded by such nucleic acids. In some embodiments LUT1 is converted to a β-hydroxylase. In some embodiments CYP97A is converted to an ε-hydroxylase. In some embodiments the location of the hydroxylation on the ring is changed (~~e.g.~~*e.g.*, from carbon 3 to carbons 2, 4, 5, ~~or 6,~~ or 6). In some embodiments, CYP97A activity is reversed to CYP97B activity. Examples of such substitutions are provided by Cunningham and Gantt E. Proc Natl Acad Sci U S A. 27;98(5):2905-10 (2001), herein incorporated by reference.

On page 55, please replace the paragraph beginning on line 3 and ending on line 8 with the following amended paragraph:

It is contemplated that *LUT1*, and in particular *lut1*, *lut1-1*, *lut1-2*, *lut1-3*, or related P450-like hydroxylases genes; genes and coding sequences (*e.g.*, any one or more of SEQ ID NOs: ~~5-9~~ 5-7, 22-27, 40-48, 53-55, and 57 ~~and 58~~ and fragments and variants thereof) can be utilized as starting nucleic acids for directed evolution. These techniques can be utilized to develop encoded LUT1 product variants having desirable properties such as increased kinetic activity or altered binding affinity.

On page 65, please replace the paragraph beginning on line 13 and ending on line 23 with the following amended paragraph:

The present invention further provides nucleic acid sequences having the coding sequence (or a portion of the coding sequence) for a LUT1 protein (*e.g.*, SEQ ID NOs: 1-4, 16-21, 33-39, and 49-52~~ and 58~~) and/or CYP97A ortholog protein fused in frame to a marker sequence that allows for expression alone or for both expression and purification of the polypeptide of the present invention. A non-limiting example of a marker sequence is a hexahistidine tag that is supplied by a vector, for example, a pQE-30 vector which adds a hexahistidine tag to the N terminal of a *LUT1* gene and/or CYP97A ortholog gene and which results in expression of the polypeptide in a bacterial host, or, for example, the marker

sequence is a hemagglutinin (HA) tag when a mammalian host is used. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein (Wilson *et al.*, Cell, 37:767 (1984), herein incorporated by reference).

On page 71, please replace the paragraph beginning on line 12 and ending on line 21 with the following amended paragraph:

In particular, some embodiments of the present invention provide recombinant constructs comprising one or more of the nucleic sequences as broadly described above (*e.g.*, SEQ ID NOs: ~~5-9~~ 5-7, 22-27, 40-48, 53-55, and 57 ~~and 58~~). In some embodiments of the present invention, the constructs comprise a vector, such as a plasmid or eukaryotic vector, or viral vector, into which a nucleic acid sequence of the invention has been inserted, in a forward or reverse orientation. Examples of such vectors of the present invention are shown in Fig. 12. In preferred embodiments of the present invention, the appropriate nucleic acid sequence is inserted into the vector using any of a variety of procedures. In general, the nucleic acid sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art.

On page 82, please replace the paragraph beginning on line 28 and ending on page 83, line 9 with the following amended paragraph:

There is no upper limit on the length of the dsRNA that can be used. For example, the dsRNA can range from about 21 base pairs (bp) of the gene to the full length of the gene or more. In one embodiment, the dsRNA used in the methods of the present invention is about 1000 bp in length. In another embodiment, the dsRNA is about 500 bp in length. In yet another embodiment, the dsRNA is about 22 bp in length. In some preferred embodiments, the sequences that mediate RNAi are from about 21 to about 23 nucleotides. That is, the isolated RNAs of the present invention mediate degradation of the target RNA (e.g., major sperm protein, chitin synthase, or RNA polymerase II). In preferred embodiments, dsRNAs corresponding to all or a portion of nucleic acids encoding a polypeptide comprising SEQ ID NOs: 1-4, 16-21, 33-39, 49-52 and 56, or nucleic acids corresponding to SEQ ID NOs: ~~5-9~~ 5-7, 22-27, 40-48, 53-55, and 57 ~~and 58~~ are utilized.

On page 85, please replace the paragraph beginning on line 5 and ending on line 29 with the following amended paragraph:

The methods of the present invention are not limited to any particular plant comprising a heterologous nucleic acid (~~e.g.~~, *e.g.*, plants comprising a heterologous nucleic acid encoding a polypeptide comprising SEQ ID NOs: 1-4, 16-21, 33-39, 49-52 and 56, or nucleic acids corresponding to SEQ ID NOs: ~~5-9~~ 5-7, 22-27, 40-48, 53-55, and 57~~and 58~~. Indeed, a variety of plants are contemplated, including but not limited to tomato, sunflowers, rice, corn, barley, wheat, *Brassica*, *Arabidopsis*, sunflower, marigolds, and soybean. The term "plant" is used in it broadest sense. It includes, but is not limited to, any species of woody, ornamental or decorative, crop or cereal, fruit or vegetable, fruit plant or vegetable plant, flower or tree, macroalga or microalga, phytoplankton and photosynthetic algae (*e.g.*, green algae *Chlamydomonas reinhardtii* and diatom *Skeletonema costatum*). It also refers to a uniclelluar plant (*e.g.* microalga) and a plurality of plant cells that are largely differentiated into a colony (*e.g.* volvox) or a structure that is present at any stage of a plant's development. Such structures include, but are not limited to, a fruit, a seed, a shoot, a stem, a leaf, a flower petal, ~~etc.~~*etc.* The term "plant tissue" includes differentiated and undifferentiated tissues of plants including those present in roots, shoots, leaves, pollen, seeds and tumors, as well as cells in culture (*e.g.*, single cells, protoplasts, embryos, callus, *etc.*). In one embodiment, transgenic seeds of the present invention may contain 5X as much β-carotene over wild-type seeds. Plant tissue may be in planta, in organ culture, tissue culture, or cell culture. The term "plant part" as used herein refers to a plant structure or a plant tissue. In some embodiments of the present invention transgenic plants are crop plants. The term "crop" or "crop plant" is used in its broadest sense. The term includes, but is not limited to, any species of plant or alga edible by humans or used as a feed for animals or fish or marine animals, or consumed by humans, or used by humans (natural pesticides), or viewed by humans (flowers) or any plant or alga used in industry or commerce or education.

On page 86, please replace the paragraph beginning on line 1 and ending on line 8 with the following amended paragraph:

The methods of the present invention contemplate the use of at least one heterologous gene encoding a *LUT1* gene, or a CYP97A gene, or encoding a sequence designed to decrease

or increase, *LUT1*, or CYP97A gene expression, as described previously ~~(e.g.~~(*e.g.*, vectors encoding a nucleic acid encoding a polypeptide comprising SEQ ID NOs: 1-4, 16-21, 33-39, 49-52 and 56, or nucleic acids corresponding to SEQ ID NOs: ~~5-9~~ 5-7, 22-27, 40-48, 53-55, and 57~~ and 58~~). Heterologous genes include but are not limited to naturally occurring coding sequences, as well variants encoding mutants, variants, truncated proteins, and fusion proteins, as described above.

On page 98, please replace the paragraphs beginning on line 24 and ending on page 99, line 12 with the following amended paragraphs:

Isolation of T-DNA Knockout Mutants in At3g53130 and Generation of a Carotenoid Hydroxylase Triple Knockout Mutant Line. At3g53130 specific primers (forward, 5'-CTTCCTCTTCTTACTCTTCTCTCTTCACT-3' (SEQ ID NO:28); reverse, 5'-AAGAACGATGGATGTTATAGACTGAAATC-3' (SEQ ID NO:29)) were sent to the University of Wisconsin Arabidopsis T-DNA knockout facility to identify knockout mutants of the *LUT1* gene. A single knockout line, designated *lut1-3*, was identified and isolated as described (http://www.biotech.wisc.edu/Arabidopsis/). In order to generate a hydroxylase triple knockout mutant line, homozygous *lut1-3* and *b1 b2* plants were crossed. Putative *lut1-3 b1 b2* triple mutants were identified from the segregating F_2 population by HPLC and their genotypes confirmed by PCR as previously described (Tian, *et al. Plant Cell* 15, 1320-1332 (2003), herein incorporated by reference).

TaqMan Real-Time PCR Assay. *LUT1* mRNA levels were quantified by TaqMan real-time PCR using elongation factor EF1a mRNA levels for normalization (Tian, *et al. Plant Cell* 15, 1320-1332 (2003), herein incorporated by reference]. The *LUT1* TaqMan probe and primers are: 5'-CCGTCTCGCTGCTGGTCCTCG-3' (SEQ ID NO:30) (TaqMan probe), 5'-GGATGAATGAGTACGGACCCAT-3' (SEQ ID NO:31) (forward primer), and 5'-GGGTCGCTCACAATTACGAAA-3' (SEQ ID NO:32) (reverse primer). The relative quantity of the transcripts was calculated using the comparative CT method [Livak, *PE applied Biosystems. User Bulletin* 2, 11-15 (1997), herein incorporated by reference].

On page 102, please replace the paragraph beginning on line 14 and ending on line 23 with the following amended paragraph:

The deduced amino acid sequence of LUT1 contains several features characteristic of cytochrome P450 enzymes (Fig. 2C). Cytochrome P450 monooxygenases contain a consensus sequence of (A/G)GX(D/E)T(T/S) (SEQ ID NO:12) that forms a binding pocket for molecular oxygen with the invariant Thr residue playing a critical role in oxygen binding in both prokaryotic and eukaryotic cytochrome P450s (Chapple, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 49, 311-343 (1998), herein incorporated by reference). In the deduced LUT1 protein sequence, this oxygen-binding pocket is highly conserved (single underlined amino acids in Fig. 2C). The conserved sequence around the heme-binding cysteine residue for cytochrome P450 type enzymes is FXXGXXXCXG (SEQ ID NO:14), and is also present in LUT1 (double underlined amino acids in Fig. 2C).

Please insert the attached Sequence Listing into the specification after the abstract.